

Analisa Kompetensi Siswa SMK N 7 Semarang Menggunakan Tehnik Datamining Metode *Neural Network*

Najib Biyantoro Hadi

*Jurusan Tehnik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Semarang, jeep.dinus@yahoo.co.id*

ABSTRAK

Ujian Nasional merupakan standarisasi dari pemerintah yang akan menunjukkan kemampuan siswa pada mata pelajaran yang diujikan, ujian nasional diadakan pada tahun terakhir belajar. Ujian nasional juga akan menentukan siswa dalam kenaikan tingkat dari SMK/SMA ke jenjang yang lebih tinggi yaitu ke bangku kuliah. Penelitian ini dilakukan pada SMK N 7 Semarang yang semua siswanya lulus, namun apakah setiap siswa mempunyai kompetensi yang diharapkan oleh pemerintah. Pada awal penelitian ini penulis mengumpulkan data nilai di SMK N 7, data nilai ujian semester dari tahun pertama hingga tahun terakhir, nilai ujian nasional di jadikan sebagai output dari penelitian ini. Penelitian menggunakan tehnik datamining dengan menggunakan metode Neural Network algoritma Backpropagation untuk memperoleh nilai rmse terkecil. Pengujian akan dikelompokkan berdasarkan mata pelajaran yaitu Matematika, Bahasa Indonesia dan Bahasa Inggris. Hasil dari pengujian diperoleh nilai rmse terkecil Matematika 3.492, Bahasa Inggris 4.268, Bahasa Indonesia 4.047 pada siswa SMK N 7 Semarang.

Keywords : datamining, rmse, Neural Network Backpropagation, Nilai Ujian

1. PENDAHULUAN

Siswa merupakan salah satu komponen proses pengajaran formal di dalam instansi pendidikan yang juga merupakan indikator kesuksesan proses pembelajaran. Dalam proses belajar siswa akan memperoleh nilai baik nilai harian, nilai ujian semester hingga ujian nasional yang diadakan tiap tahun untuk menentukan kelulusan siswa. Nilai tersebut akan di simpan dalam arsip baik secara komputasi atau dengan catatan kertas.

Penerapan sistem informasi pada suatu instansi merupakan kebutuhan pokok, namun seiring dengan berjalannya waktu penggunaan sistem informasi ini mengakibatkan penumpukan data. Karena data tersebut maka munculah tehnik data mining yang akan mengolah kembali data yang

menumpuk kemudian dapat diambil pola atau gambaran dari tumpukan data.

Proses *data mining* sendiri adalah Kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola dan hubungan dalam set data berukuran besar (*Santosa, 2007*). Perkembangan data mining yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar. Namun demikian, pertumbuhan yang pesat dari akumulasi data itu telah menciptakan kondisi yang sering disebut sebagai "*Rich of Data but Poor of Information*" karena data yang terkumpul itu tidak dapat digunakan untuk aplikasi yang berguna. Tidak jarang kumpulan data itu dibiarkan begitu saja seakan-akan "kuburan data" (*data tombs*) (*Pramudiono, 2003:1*).

Penggunaan metode neural network pada penelitian ini karena neural network mempunyai kemampuan untuk menyelesaikan masalah yang rumit, pengenalan pola, prediksi pasar saham dan sebagainya. Berdasarkan analisa diatas data yang diperoleh dari SMK N 7 mempunyai tipe numerik maka akan digunakan metode neural network struktur *backpropagation*.

Penelitian dilakukan dengan menggunakan metode CRISP-DM yang mempunyai beberapa langkah seperti dibawah ini :

1. Pengumpulan data
Dalam tahap ini terbagi menjadi dua bagian yaitu pemahaman bisnis dan pemahaman data, untuk mendapat data diatas dilakukan observasi pada obyek penelitian.
2. Pengolahan data
Setelah memperoleh data maka akan dipilih data yang sesuai kebutuhan, di SMK data yang digunakan yaitu nilai untuk ujian nasional yakni Matematika, Bahasa Indonesia dan Bahasa Inggris.
3. Pemodelan (*Modelling*)
Pada tahap ini dilakukan pemilihan model apa yang akan digunakan untuk mengolah data. berdasarkan data yang diperoleh maka model yang dipakai *Neural Network backpropagation*.
4. Validasi dan evaluasi
Sesuai dengan metode yang digunakan validasi yang dilakukan menggunakan *rmse (root mean square error)*.
5. *Deployment*
Penyebaran hasil penelitian ini akan digunakan oleh pihak sekolah untuk mengetahui kompetensi siswa berdasarkan nilai *rmse*.

2. NEURAL NETWORK DAN BACK PROPAGATION

2.1 NEURAL NETWORK

Neural Network atau dalam bahasa Indonesia disebut Jaringan Saraf

Tiruan menggunakan ide dari jaringan saraf manusi. Algoritma NN di temukan pertama kali pada tahun 1943 oleh neurofisiologi Warren McCulloch dan Walter Pits ahli logika. NN merupakan sebuah teknologi komputasi, tidak memberikan suatu keajaiban tetapi jika digunakan dengan tepat akan menghasilkan suatu hasil yang luar biasa. Kemampuan NN untuk menyelesaikan masalah yang rumit telah dibuktikan dalam berbagai macam penelitian seperti analisa data, meteorologi, pengenalan pola, sistem kontrol, deteksi fenomena kedokteran, prediksi pasar saham, dan sebagainya (Yani, 2005). Berikut ini beberapa pengertian *neural network* dari berbagai sumber:

1. Hecht-Nielsend (1988 "Suatu neural network (NN), adalah suatu struktur pemroses informasi yang terdistribusi dan bekerja secara paralel, yang terdiri atas elemen pemroses (yang memiliki memori lokal dan beroperasi dengan informasi lokal) yang diinterkoneksi bersama dengan alur sinyal searah yang disebut koneksi. Setiap elemen pemroses memiliki koneksi keluaran tunggal yang bercabang (fan out) ke sejumlah koneksi kolateral yang diinginkan (setiap koneksi membawa sinyal yang sama dari keluaran elemen pemroses tersebut). Keluaran dari elemen pemroses tersebut dapat merupakan sebarang jenis persamaan matematis yang diinginkan. Seluruh proses yang berlangsung pada setiap elemen pemroses harus benar-benar dilakukan secara lokal, yaitu keluaran hanya bergantung pada nilai masukan pada saat itu yang diperoleh melalui koneksi dan nilai yang tersimpan dalam memori lokal".
2. Menurut Haykin (1994) mendefinisikan jaringan saraf sebagai berikut: Sebuah jaringan

saraf adalah sebuah prosesor yang terdistribusi paralel dan mempunyai kecenderungan untuk menyimpan pengetahuan yang didapatkannya dari pengalaman dan membuatnya tetap tersedia untuk digunakan. Hal ini menyerupai kerja otak dalam dua hal yaitu: 1. Pengetahuan diperoleh oleh jaringan melalui suatu proses belajar. 2. Kekuatan hubungan antar sel saraf yang dikenal dengan bobot sinapsis digunakan untuk menyimpan pengetahuan.

3. Menurut Zurada (1992), sistem saraf tiruan atau jaringan saraf tiruan adalah sistem selular fisik yang dapat memperoleh, menyimpan dan menggunakan pengetahuan yang didapatkan dari pengalaman.

Karakteristik yang dimiliki *neural network* (Fausett, 1994) seperti berikut :

1. Memiliki kemampuan menghasilkan output terhadap pola yang belum pernah dipelajari (*generalization*).
2. Mempunyai kemampuan untuk memproses inputan yang terdapat kesalahan didalamnya dengan tingkat toleransi tertentu.
3. Mampu beradaptasi dengan perubahan yang terjadi terhadap nilai input dan output. Bentuk adaptasi ini umumnya diwujudkan dalam perubahan nilai bobot.
4. Akurasi prediksi pada pada umumnya cukup tinggi dan memerlukan waktu yang relatif lama dalam pembelajaran.

2.2 ALGORITMA BACKPROPAGATION

Diperkenalkan oleh G.E Hinton, E. Rumelhart dan R.J Williams pada tahun 1986, menurut Daghli (1994) jenis *Backpropagation* lebih flaksibel dan relatif lebih baik karena banyak diterapkan dalam khususnya untuk dunia industri. Algoritma BP biasanya digunakan untuk preceptron berlapis banyak,

preceptron paling tidak mempunyai input, output dan beberapa output dengan lapisan tengah disebut hidden layer atau dikenal dengan sebutan layer tersembunyi, bisa satu atau lebih hidden layer tetapi dalam prakteknya paling banyak menggunakan tiga lapis hidden layer. *Output* terakhir dari *Hidden Layer* langsung dipakai untuk sebagai *output* dari NN (Hagan, 1996).

Proses pelatihan *Backpropagation* memerlukan tiga tahapan yaitu *feedforward* data input untuk pelatihan, *backpropagation* untuk nilai *error* dan penyesuaian nilai bobot tiap *node* masing-masing *layer* NN.

Feedforward nilai input tiap unit ke-*i* (x_i) menerima sinyal input yang akan dilanjutkan ke *hidden layer* z_1, z_2, \dots, z_i selanjutnya *hidden* unit ke-*j* akan menghitung nilai sinyal (z_j) yang akan dipancarkan ke *output layer* menggunakan fungsi aktivasi f .

$$z_{in_j} = \phi_{1j} + \sum_{i=1}^n x_i v_j$$

Dan

$$z_j = f(z_j)$$

dimana ϕ_{1j} = bias *hidden* unit ke-*j*. Nilai bias dan bobot awal dapat diambil secara acak. Tiap unit *output* ke-*k* (y_k).

$$y_k = \phi_{1k} + \sum_{i=1}^n x_i v_j$$

Dan

$$y_k = f(y_k)$$

3. Fungsi Aktifasi

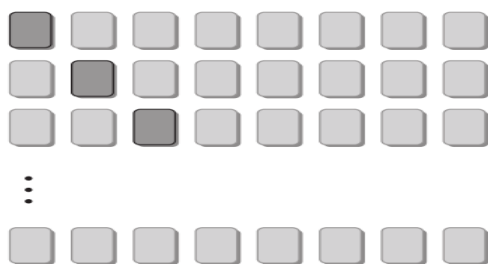
Karakteristik yang harus dimiliki fungsi aktivasi adalah kontinyu, differensiabel dan monoton menurun. Salah satu fungsi yang sering digunakan adalah fungsi signoid yang memiliki range (0,1)

$$f(x) = \frac{1}{1 + \exp(-x)}$$

node pada *output* layer memiliki nilai antara 0-1. Dari beberapa pemaparan fungsi aktivasi diatas fungsi aktivasi *sigmoid biner* yang digunakan untuk algoritma backpropagation.

4. VALIDASI

Cross validation adalah teknik pengambilan sampel secara random yang menjamin setiap jumlah kemunculan data yang diamati sama dengan jumlah data *training* dan hanya sekali pada data *testing* (Vercellis, 2009). Dalam *cross validation* kita harus menetapkan jumlah partisi atau *fold*, standar yang biasa digunakan untuk memperoleh estimasi kesalahan terbaik adalah 10 kali partisi atau *tenfold cross-validation* (Gorunescu, 2011). Data dibagi secara random menjadi 10 bagian dengan perbandingan yang sama kemudian *error rate* dihitung bagian demi bagian, selanjutnya *error rate* secara keseluruhan diperoleh dari menghitung rata-rata *error rate* dari 10 bagian.



Gambar 1 Ilustrasi *tenfold cross validation*

5. HASIL DAN PENGUJIAN

5.1 Mata Pelajaran Matematika

Pertama akan dilakukan percobaan pada mata pelajaran matematika. Berikut ini adalah hasil dari

percobaan yang telah dilakukan untuk menentukan *training cycle* :

Tabel 1 Percobaan *training cycle*
Matematika

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.3	0.2	3.704
200	0.3	0.2	4.105
300	0.3	0.2	4.241
400	0.3	0.2	4.794
500	0.3	0.2	5.045
600	0.3	0.2	5.522
700	0.3	0.2	5.683
800	0.3	0.2	5.732
900	0.3	0.2	5.736
1000	0.3	0.2	5.694

Nilai *training cycle* yang dipilih berdasarkan nilai *root mean square error (rmse)* yang terkecil. Berdasarkan hasil diatas maka nilai *rmse* yang terkecil adalah 100 *training cycle* Untuk selanjutnya akan digunakan untuk menentukan nilai *learning rate*. Untuk memperoleh nilai *learning rate* dilakukan percobaan dengan menggunakan nilai *training cycle* 100 dan nilai *learning rate* antara 0.1 sampai dengan 0.9. Sedangkan nilai momentum yang digunakan adalah nilai 0.2, berikut ini adalah hasil dari percobaan penentuan *learning rate* :

Tabel 2 Percobaan *learning rate*
Matematika

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.1	0.2	3.505
100	0.2	0.2	3.589
100	0.3	0.2	3.704
100	0.4	0.2	3.833
100	0.5	0.2	4.227
100	0.6	0.2	4.370
100	0.7	0.2	5.198
100	0.8	0.2	5.712
100	0.9	0.2	5.987

Nilai *learning rate* yang menghasilkan nilai *rmse* yang terkecil adalah 0.1 berdasarkan percobaan diatas akan dilakukan percobaan selanjutnya yaitu untuk menentukan nilai momentum. Nilai momentum diperoleh dengan memasukkan nilai 0 hingga 0.9. berikut ini adalah hasil dari percobaan penentuan momentum:

Tabel 3 Percobaan *momentum*

Matematika

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.1	0.1	3.502
100	0.1	0.2	3.505
100	0.1	0.3	3.500
100	0.1	0.4	3.492
100	0.1	0.5	3.499
100	0.1	0.6	3.549
100	0.1	0.7	3.625
100	0.1	0.8	3.983
100	0.1	0.9	6.012

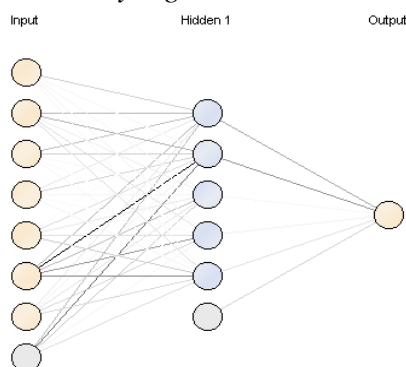
Berdasarkan percobaan diatas maka untuk parameter *neural network* yang dipilih adalah 0.4 pada mata pelajaran matematika. Penentuan hidden kayer dilakukan 1-14, berikut ini hasil percobaan yang telah dilakukan :

Table 4 penentuan *hidden layer*

<i>Hidden layer</i>	RMSE
1	3.559
2	3.533
3	3.538
4	3.899
5	3.492
6	3.584
7	3.799
8	3.544
9	3.535
10	3.583

11	3.791
12	3.419
13	3.580
14	3.576

Dari percobaan diatas maka didapat parameter yang menghasilkan nilai *rmse* terkecil yaitu *training cycle* 100, *learning rate* 0.1, *momentum* 0.4 dan 5 *hidden layer*.

Gambar 2 *Neural network yang terbentuk*

5.2 Mata pelajaran Bahasa Indonesia

Percobaan selanjutnya pada mata pelajaran Bahasa Indonesia. Percobaan untuk menentukan *training cycle* :

Tabel 5 Percobaan *training cycle*

bahasa Indonesia

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.3	0.2	4.741
200	0.3	0.2	4.901
300	0.3	0.2	4.635
400	0.3	0.2	4.549
500	0.3	0.2	4.620
600	0.3	0.2	5.004
700	0.3	0.2	5.578
800	0.3	0.2	6.043
900	0.3	0.2	6.240
1000	0.3	0.2	6.323

Nilai *training cycle* yang diperoleh adalah 400, yang menghasilkan nilai *rmse* terkecil, seperti pada percobaan sebelumnya maka akan dilanjutkan melakukan percobaan untuk mencari nilai *learning rate* dengan range dari 0.1 sampai 0.7, percobaan hanya dilakukan sampai 0.7 dikarenakan nilai yang didapat jauh melebihi dari *learning rate* sebelumnya. Berikut ini adalah percobaan untuk mencari nilai *learning rate*:

Tabel 6 Percobaan Penentuan *learning rate*

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
400	0.1	0.2	4.556
400	0.2	0.2	4.371
400	0.3	0.2	4.549
400	0.4	0.2	5.284
400	0.5	0.2	6.389
400	0.6	0.2	6.884
400	0.7	0.2	11.584

Berdasarkan percobaan diatas nilai *learning rate* yang didapat adalah 0.2, kemudian digunakan untuk mencari nilai momentum dengan range antara 0.1 hingga 0.9, berikut ini adalah hasil dari percobaan penentuan momentum:

Tabel 7 Percobaan Penentuan momentum

TAININ G CYCLE	LEARNI NG RATE	MOMONT UM	RMSE
400	0.2	0.1	4.350
400	0.2	0.2	4.371
400	0.2	0.3	4.412
400	0.2	0.4	4.672
400	0.2	0.5	4.629
400	0.2	0.6	5.349

400	0.2	0.7	6.876
400	0.2	0.8	9.545
400	0.2	0.9	10.108

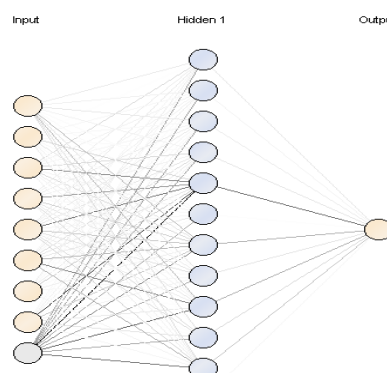
Seperti pada tabel percobaan diatas maka nilai momentum, yang diambil adalah 0.1 yang menghasilkan nilai *rmse* 4.350. selanjutnya penentuan hidden layer seperti tabel dibawah ini :

Table 8 penentuan *hidden layer*

<i>Hidden layer</i>	RMSE
1	4.661
2	4.440
3	4.857
4	4.849
5	4.680
6	4.350
7	4.506
8	4.520
9	4.503
10	4.881
11	4.047
12	4.746
13	4.489
14	4.086

Setelah percobaan diatas maka nilai yang diinputkan dalam parameter *neural network* dengan nilai 400 pada *training cycle*, 0.1 pada nilai *learning rate*, 0.2 momentum dan 11 *hidden layer*.

Gambar 3 *neural network* yang terbentuk



5.3 Mata Pelajaran Bahasa Inggris

Yang terakhir akan dilakukan percobaan pada mata pelajaran bahasa inggris. Seperti pada percobaan sebelumnya pertama kali akan dilakukan pencarian *training cycle* terbaik untuk memperoleh nilai *rmse* terkecil seperti tael dibawah ini:

Tabel 9 Percobaan Penentuan *training cycle*

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.3	0.2	5.642
200	0.3	0.2	5.773
300	0.3	0.2	6.326
400	0.3	0.2	7.035
500	0.3	0.2	7.770
600	0.3	0.2	8.728
700	0.3	0.2	8.844
800	0.3	0.2	9.169
900	0.3	0.2	9.392
1000	0.3	0.2	9.550

Seperti yang terlihat pada tabel percobaan diatas nilai *training cycle* yang diambil adalah 100, kemudian percobaan untuk mencari nilai *learning rate* seperti tabel dibawah ini:

Tabel 10 Percobaan Penentuan *learning rate*

TAINING CYCLE	LEARNING RATE	MOMONTUM	RMSE
100	0.1	0.2	4.872
100	0.2	0.2	5.148
100	0.3	0.2	5.462
100	0.4	0.2	5.895
100	0.5	0.2	6.566
100	0.6	0.2	6.745
100	0.7	0.2	7.336

100	0.8	0.2	7.537
-----	-----	-----	-------

Seperti yang terlihat pada tabel diatas nilai *learning rate* yang diperoleh adalah 0.1 yang menghasilkan nilai *rmse* yang terkecil. Kemudian langkah selanjutnya adalah pencarian nilai momentum, berikut ini adalah hasil dari percobaan penentuan momentum:

Tabel 11 Percobaan Penentuan momentum

TAININ G CYCLE	LEARNI NG RATE	MOMONT UM	RMS E
100	0.1	0.1	4.857
100	0.1	0.2	4.872
100	0.1	0.3	4.896
100	0.1	0.4	4.939
100	0.1	0.5	5.003
100	0.1	0.6	5.085
100	0.1	0.7	5.173
100	0.1	0.8	5.678
100	0.1	0.9	7.184

Berdasarkan percobaan diatas maka untuk parameter *neural network* untuk nilai momentum yang dipilih adalah 0.1 pada mata pelajaran bahasa inggris.

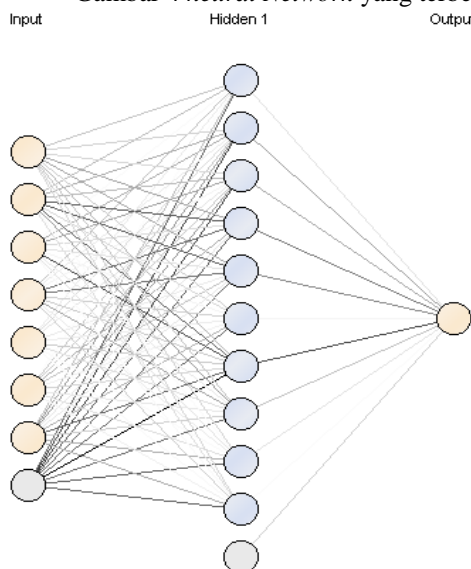
Tabel 12 Percobaan *hidden layer*

<i>Hidden layer</i>	RMSE
1	4.630
2	4.509
3	4.794
4	4.739
5	4.857
6	4.700
7	4.338
8	4.504
9	4.715
10	4.268

11	4.621
12	5.036
13	4.704
14	4.638

Hidden layer yang diperoleh adalah 10 *hidden layer*, setelah melakukan percobaan diatas maka di peroleh nilai parameter terbaik untuk memperoleh nilai *rmse* terkecil yaitu dengan training cycle 100, learning rate 0.1, *momentum* 0.1 dan *hidden layer* 10.

Gambar 4 *neural Network* yang terbentuk



6. KESIMPULAN DAN SARAN

Pada penelitian ini dilakukan pemodelan dengan menggunakan *neural network backpropagation* pada SMK N 7 Semarang, percobaan dilakukan pada mata pelajaran yang diujikan dalam Ujian Nasional, nilai *rmse* yang terkecil menunjukkan model yang lebih bagus. Hasil yang didapat dari penelitian diatas adalah :

Tabel 5.1 Hasil penelitian

Mapel	Learning rate	Training cycle	Momentum	Hidden layer	Rmse
Matematika	0.1	100	0.4	5	3.492
Bahasa	0.2	400	0.1	11	4.047

Indonesia					
Bahasa Inggris	0.1	100	0.1	10	4.268

Dari tabel diatas dapat dilihat bahwa mata pelajaran bahasa Inggris mempunyai nilai *error* terbesar, dapat diartikan bahwa siswa kurang menguasai mata pelajaran bahasa Inggris.

Saran

Dalam penelitian ini masih banyak kekurangan dikarenakan keterbatasan waktu dan sumber daya, untuk penelitian selanjutnya terdapat beberapa saran sebagai berikut :

1. Penggunaan data yang lebih variatif agar data yang dihasilkan lebih akurat.
2. Menggunakan lebih banyak *mining tool* dan menggabungkan algoritma lain untuk memperoleh hasil yg lebih baik.

DAFTAR PUSTAKA

- [1] Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3, Tahun 2013.
- [2] Susanto Ph.D, Sani dan Dedy Suryadi, ST.,MS, "PENGANTAR DATA MINING", Andi Yogyakarta : 2010.
- [3] Yuli Fitrianto, "PENERAPAN NEURAL NETWORK UNTUK PREDIKSI DATA TIME-SERIES ARUS LALU LINTAS JANGKA PENDEK", Thesis UNIVERSITAS DIAN NUSWANTORO : 2011.
- [4] Larose, D.T. 2005. "Discovering knowledge in data : An Introduction to data mining". Hoboken: Willey-Interscience, John Willey and Sons, Inc.
- [5] Imron Rosyidi, Mochamad Hariadi dan I ketut Eddy Purnama, "DATA MINING KEMAMPUAN SISWA BERBASIS NEURO FUZZY" Jurnal SESINDO ITS: 2011.
- [6] Novi Yanti, "PENERAPAN METODE NEURAL NETWORK DENGAN STRUKTUR BACKPROPAGATION

UNTUK PREDIKSI STOK TOKO OBAT DI APOTIK (STUDI KASUS : APOTEK ABC)". Seminar Nasional Aplikasi Teknologi Informasi 2011 (SNATI2011).

- [7] Imam Eko Ardianto, "MODEL ESTIMASI *NEURAL NETWORK*, APLKASI PERAMALAN TINGKAT BAGI HASIL DEPOSITO *MUDHARABAH* DENGAN VARIABLE MAKROEKONOMI SEBAGAI PENENTU", SKRIPSI. UNIVERSITAS DIPONEGORO:2012.
- [8] Fausett, Laurene V. 1994. Fundamentals of Neural Networks:Architectures, Algorithms, and Applications".